JB Journals.ASM.org

GENOME ANNOUNCEMENT

Genome Sequence of the Human- and Animal-Pathogenic Strain *Nocardia cyriacigeorgica* GUH-2

Anthony Zoropogui,[a] Petar Pujic,[b] Philippe Normand,[b] Valérie Barbe,[c] Blaine Beaman,[d] LoVelle Beaman,[d] Patrick Boiron,[a] Céline Colinon,[a] Amélie Deredjian,[a] Arnault Graindorge,[a] Sophie Mangenot,[c] Sylvie Nazaret,[b] Manuelle Neto,[a] Stéphanie Petit,[a] David Roche,[c] David Vallenet,[c] Veronica Rodríguez-Nava,[a] Yves Richard,[a] Benoit Cournoyer,[a] and Didier Blaha[a]

Université de Lyon, Research Group on Bacterial Opportunistic Pathogens and Environment, Biological Resource Center, UMR5557 Ecologie Microbienne, Université Lyon 1, CNRS, and VetAgro Sup Veterinary Campus, Lyon, France[a]; Université de Lyon 1, Université de Lyon, CNRS, Ecologie Microbienne UMR5557, Villeurbanne, France[b]; CEA/IG/Genoscope, Evry, France[c]; and Department of Medical Microbiology and Immunology, University of California, Davis, Davis, California, USA[d]

The pathogenic strain *Nocardia cyriacigeorgica* GUH-2 was isolated from a fatal human nocardiosis case, and its genome was sequenced. The complete genomic sequence of this strain contains 6,194,645 bp, an average G+C content of 68.37%, and no plasmids. We also identified several protein-coding genes to which *N. cyriacigeorgica*'s virulence can potentially be attributed.

Nocardia cyriacigeorgica is a species representing the filamentous actinobacteria, which are frequently isolated from diseased human and animal tissues (3, 13). This infection, known as nocardiosis, has been found in the lung, brain, heart, kidney, skin, and eye tissues (1, 5, 10). Immune-compromised individuals are the most vulnerable to infection; however, healthy individuals can also be affected (7). *N. cyriacigeorgica* is the most prevalent species responsible for nocardiosis in North America (9, 12), while placing fourth in France (11). The strain *N. cyriacigeorgica* GUH-2 was isolated at Washington, DC, in the 1970s (3) and is a model organism for studying nocardial infection. We used a hybrid Sanger/pyrosequencing approach to determine its genomic sequence. Two shotgun libraries (3 kb and 10 kb) were constructed, and 82,125 (3 kb) and 24,217 (10 kb) Sanger reads were produced using ABI 3730 Applera (Applied Biosystems) machines. This was complemented by 837,088 reads with FLX genome sequencers (Roche Applied Science). The Arachne "HybridAssemble" software program (Broad Institute [www.broad.mit.edu]) was used for hybrid assembly. The assembly validation was done using the Mekano interface (Genoscope). Genome finishing was achieved by primer walking using the *in vitro* transposition technology (Template Generation System II kit; Finnzyme, Espoo, Finland). Coding sequences (CDSs) were predicted using the AMIGene software program (4) and the automated functional annotation platform Microscope (14). Manual scrutiny of the automatic annotation was performed using the Web-based MaGe (Magnifying Genomes) interface (15). In total, 5,491 predicted gene functions were manually scrutinized.

The genome of *N. cyriacigeorgica* GUH-2 is a circular chromosome of 6,194,645 bp and has a G+C content of 68.37%. This genome encodes 49 tRNA genes, 3 rRNA operons, and 5,491 protein CDSs. Fourteen CDSs were found to contain frameshifts. Functions were assigned to 62.12% of the CDSs. Of this percentage, 28.11% encode conserved proteins of unknown function and 9.71% of them had no homology to known proteins.

For several CDSs, involvement in virulence can be suggested on the basis of similarity to *Mycobacterium tuberculosis* and *Nocardia farcinica* virulence genes (6, 8). These genes include six complete *mce* loci that are used for mammalian cell entry; 17 esterases, including 3 85-kDa-antigen family proteins; 19 lipoproteins; hemolysin; and 5 PE PGRS/PPE (Pro-Glu proteins containing polymorphic GC-rich sequences, Pro-Pro-Glu) family proteins. Furthermore, two superoxide dismutases and three catalase genes were also found to be similar to genes involved in macrophage resistance in *Nocardia* (2).

Nucleotide sequence accession number. The genome of *N. cyriacigeorgica* GUH-2 is available at the EMBL under the accession number FO082843.

ACKNOWLEDGMENTS

This project, no. 110/AP2006-2007, was carried out in collaboration with Platforms Génomique Environmentale UMR CNRS-UCBL 5557 and Genoscope. A.Z. was supported by a Ph.D. fellowship from the Cluster Environnement program of the Rhône-Alpes Region, France, and the University Lyon-1. This work was partially funded by the CNRS (Programme Interdisciplinaire CNRS, Maladies Infectieuses et Environnement), IMU-LABEX (Intelligence of the Urban Worlds), University Lyon-1, Rhône-Alpes Region, and the French Research Funding Agency, ANR, SEST 2005 program no. 0005605.

REFERENCES

1. **Akcaglar S, et al.** 2008. *Nocardia cyriacigeorgica*: pulmonary infection in a patient with Basedow-Graves disease and a short review of reported cases. Int. J. Infect. Dis. **12**:335–338.
2. **Beaman BL, Black CM, Doughty F, Beaman L.** 1985. Role of superoxide dismutase and catalase as determinants of pathogenicity of *Nocardia asteroides*: importance in resistance to microbicidal activities of human polymorphonuclear neutrophils. Infect. Immun. **47**:135–141.
3. **Beaman BL, Maslan S.** 1977. Effect of cyclophosphamide on experimental *Nocardia asteroides* infection in mice. Infect. Immun. **16**:995–1004.
4. **Bocs S, Cruveiller S, Vallenet D, Nuel G, Medigue C.** 2003. AMIGene: Annotation of MIcrobial Genes. Nucleic Acids Res. **31**:3723–3726.
5. **Cargill JS, Boyd GJ, Weightman NC.** 2010. *Nocardia cyriacigeorgica*: a case of endocarditis with disseminated soft-tissue infection. J. Med. Microbiol. **59**:224–230.
6. **Cole ST, et al.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393**:537–544.

Received 2 February 2012 **Accepted** 9 February 2012

Address correspondence to Petar Pujic, petar.pujic@univ-lyon1.fr.

doi:10.1128/JB.00161-12

7. **Green JS, Abeles SR, Uslan DZ, Mehta SR.** 2011. Persistent neutrophilic meningitis in an immunocompetent patient after basilar skull fracture: case report. BMC Infect. Dis. **11**:136.
8. **Ishikawa J, et al.** 2004. The complete genomic sequence of *Nocardia farcinica* IFM 10152. Proc. Natl. Acad. Sci. U. S. A. **101**:14925–14930.
9. **McTaggart LR, Richardson SE, Witkowska M, Zhang SX.** 2010. Phylogeny and identification of *Nocardia* species on the basis of multilocus sequence analysis. J. Clin. Microbiol. **48**:4525–4533.
10. **Pujic P, Beaman BL.** 2001. Actinomyces and Nocardia. *In* Max Sussman (ed), Molecular medical microbiology. Academic Press, San Diego, CA.
11. **Rodriguez-Nava V, et al.** 2008. Nocardiosis, an expanding disease. Antibiotiques **10**:115–127.
12. **Schlaberg R, Huard RC, Della-Latta P.** 2008. *Nocardia cyriacigeorgica*, an emerging pathogen in the United States. J. Clin. Microbiol. **46**:265–273.
13. **St Leger JA, et al.** 2009. Comparative pathology of nocardiosis in marine mammals. Vet. Pathol. **46**:299–308.
14. **Vallenet D, et al.** 2009. MicroScope: a platform for microbial genome annotation and comparative genomics. Database (Oxford) **2009**:bap021.
15. **Vallenet D, et al.** 2006. MaGe: a microbial genome annotation system supported by synteny results. Nucleic Acids Res. **34**:53–65.